



## Comparison of Simulated Spatial Patterns Using Rain Gauge and Polarimetric-Radar-Based Precipitation Data in Catchment Hydrological Modeling

He, Xin; Koch, Julian; Zheng, Chunmiao; Bovith, Thomas; Jensen, Karsten H.

*Published in:*  
Journal of Hydrometeorology

*DOI:*  
[10.1175/JHM-D-17-0235.1](https://doi.org/10.1175/JHM-D-17-0235.1)

*Publication date:*  
2018

*Document version*  
Publisher's PDF, also known as Version of record

*Document license:*  
[Unspecified](#)

*Citation for published version (APA):*  
He, X., Koch, J., Zheng, C., Bovith, T., & Jensen, K. H. (2018). Comparison of Simulated Spatial Patterns Using Rain Gauge and Polarimetric-Radar-Based Precipitation Data in Catchment Hydrological Modeling. *Journal of Hydrometeorology*, 19(8), 1273-1288. <https://doi.org/10.1175/JHM-D-17-0235.1>

# Comparison of Simulated Spatial Patterns Using Rain Gauge and Polarimetric-Radar-Based Precipitation Data in Catchment Hydrological Modeling

XIN HE

*School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China,  
and Department of Hydrology, Geological Survey of Denmark and Greenland, Copenhagen, Denmark*

JULIAN KOCH

*Department of Hydrology, Geological Survey of Denmark and Greenland, Copenhagen, Denmark*

CHUNMIAO ZHENG

*School of Environmental Science and Engineering, Southern University of Science and Technology, and Guangdong  
Provincial Key Laboratory of Soil and Groundwater Pollution Control, Shenzhen, China*

THOMAS BØVITH

*Department of Research and Development, Danish Meteorological Institute, Copenhagen, Denmark*

KARSTEN H. JENSEN

*Department of Geosciences and Natural Resource Management, University of Copenhagen, Copenhagen, Denmark*

(Manuscript received 4 December 2017, in final form 27 April 2018)


## ABSTRACT

With the advance of the weather radar technology, dual-polarization (dual-pol) radar data are now available for hydrological studies, which go beyond the traditional rainfall products relying purely on rain gauge data. Previous studies have focused on the evaluation of rainfall products and their hydrological responses using point-based observational data; however, spatial patterns of simulated hydrological variables are equally important to be considered in order to fully address the distributed effect of the precipitation estimates. In the present study, we compare three rainfall estimations based on rain gauge, single-polarization, and dual-pol radar data. Special attention is given to the use of the two radar products and their corresponding hydrological simulations of both surface water and groundwater. Performance of the hydrological simulations is evaluated based first on traditional point-based observations of stream discharge and groundwater head, and second on remotely sensed land surface temperature data. For the latter, the empirical orthogonal function analysis, which quantifies spatial pattern similarities, is employed. The Skjern River catchment in western Denmark is selected as the study site, and the results show that all three models perform equally well in terms of the traditional aggregated evaluation criteria, such as Nash–Sutcliffe efficiency (NSE) and RMSE on time series data. It is found that the differences of simulated hydrological spatial patterns are sensitive to rainfall signal intensity, as well as the simulation scale in space ( $<100\text{ km}^2$ ) and time (subdaily). Our study suggests that the currently available observational data have limited capabilities to clearly differentiate the performance of the three applied models due to the low resolution.

## 1. Introduction

Precipitation is the driving force of the majority of the land surface and subsurface hydrological processes, and it is therefore of critical importance in catchment hydrology. Estimation of precipitation in an accurate and meaningful way is a key element when trying to close the water budget at catchment and subcatchment

---

 Denotes content that is immediately available upon publication as open access.

---

*Corresponding author:* Xin He, hex@sustc.edu.cn

scales using hydrological models (Kuczera and Williams 1992; Boyle et al. 2001; Younger et al. 2009; Rice et al. 2015). Numerous studies have attempted to estimate precipitation quantitatively for the purpose of hydrological modeling, especially for distributed models, and these studies generally involve two aspects: *areal mean* and *spatial pattern* (Yang et al. 2010; Villarini et al. 2011; He et al. 2013; Maussion et al. 2014). Areal mean refers to the averaged amount of precipitation across the entire catchment, for example, averaging of rain gauge observational data. Spatial pattern, on the other hand, focuses on the internal distribution of precipitation within the catchment, which can be estimated by interpolation of point data or range scanning.

Weather radar is range-scanning equipment that can be used to remotely estimate both the intensity and spatial pattern of precipitation. Application of weather-radar-based quantitative precipitation estimation (QPE) in catchment hydrology has gained increasing popularity due to its high spatial and temporal resolution as well as its large spatial coverage in a fully automated way (Carpenter et al. 1999; Cole and Moore 2008, 2009; He et al. 2013; Goudenhoofd and Delobbe 2016). However, since QPE products based on radar data alone are an indirect proxy of rainfall that falls on the ground surface, they thus often bear a high degree of uncertainty. Consequently, a combination of QPE products and rain gauge data is often applied to achieve better-quality rainfall estimation (Dong et al. 2005; Bárdossy and Das 2008; Arsenault and Brisette 2014). In such frameworks, rain gauge data are mainly responsible for obtaining the correct areal mean, whereas radar data have the advantage of estimating the spatial pattern.

Recent studies have suggested that uncertainty in radar-estimated rainfall can be further addressed and the quality of QPE products can potentially be improved moving from traditional single-polarization (single-pol) technology to polarimetric [dual-polarization (dual-pol)] technology (Berne and Krajewski 2013). Dual-pol radars are able to measure not only the signal strength of reflectivity but also deduce the type, shape, size distribution, and fall behavior of the hydrometeors and therefore have advantages over single-pol radars when it comes to QPE. Various algorithms for retrieval of rainfall estimates using dual-pol data exist, for example, based on differential reflectivity, differential phase, or the combination of several radar parameters (Ryzhkov et al. 1997, 2005; Cifelli et al. 2011; Ryzhkov et al. 2014; Chang et al. 2016). Using dual-pol-based QPE product for surface water modeling has been investigated in a number of studies (Gourley et al. 2010; Cunha et al. 2013; Gao et al. 2016). However, the difference between

single-pol- and dual-pol-based QPE products for estimating rainfall spatial patterns has not yet been demonstrated. Moreover, the difference between the two types of radar QPE products for long-term (multi-annual) continuous hydrological simulation for both surface water and groundwater has not been studied. Both aspects are addressed in this study to guide the community on the differences and implications of the two different QPE products.

Suitable spatial performance measures are required for a meaningful quantification of the impact on simulated spatial patterns in a catchment model forced by the two radar products. The necessity to apply adequate statistical measures to quantify spatial similarity between patterns has frequently been demanded in order to overcome possible limitations associated with a simple cell-to-cell-based comparison that may omit pattern information (Grayson et al. 2002; Wealands et al. 2005; Koch et al. 2017). Moreover, in order to assess the performance of distributed hydrological variables generated by a distributed hydrological model, one has to take spatial observations into consideration, because temporal observations, such as streamflow, are found insensitive to differentiate the simulated spatial variability (Clark et al. 2011; Stisen et al. 2011; Koch et al. 2016a). Such observations are broadly facilitated by remote sensing products where information contained in the thermal bands is used to retrieve land surface temperature (LST), which is a variable closely related to the energy and water cycle at the land surface (Orth et al. 2017). Previously, spatial-pattern-oriented model evaluation has been successfully applied to diagnose spatial model deficiencies that would remain undetected when evaluating the model performance against traditional streamflow data (Immerzeel and Droogers 2008; Schuurmans et al. 2011; Mendiguren et al. 2017).

The need to quantitatively compare spatial patterns is prevalent throughout the Earth sciences, which has resulted in numerous pattern comparison algorithms (Roberts and Lean 2008; Li et al. 2009; Renard and Allard 2013). Among them, one promising approach is the empirical orthogonal function (EOF) analysis, which is used to quantify the spatial similarity between spatial patterns and their temporal dynamics. The approach has been successfully applied to spatially validate distributed hydrological models at various scales (Fang et al. 2015; Koch et al. 2016b; Ruiz-Pérez et al. 2017). Rainfall estimates based on weather radar have distinctive spatial patterns, and the proposed EOF analysis may serve as a good indicator to evaluate the difference between QPE products. To our knowledge, it is the first

comparison of single-pol and dual-pol radar rainfall estimation with a focus on their spatial patterns as well as their performances in hydrological simulations using a spatial-pattern-oriented metric.

The objectives of the present study are 1) to compare each rainfall product based on rain gauge, single-pol, and dual-pol radar and identify the temporal scale at which they differ; 2) to validate the hydrological simulations and quantify the model performance using different model forcings; and 3) to compare the spatial patterns of the hydrological simulations and identify the spatial scale at which they differ.

## 2. Materials and methods

### a. Weather radar and data

Polar volume weather radar data from a C-band, dual-pol weather radar operated by the Danish Meteorological Institute (DMI) was available for this study. The weather radar manufactured by Enterprise Electronics Corporation (EEC) and installed in 2008 is situated in Virring in Jutland, Denmark, at 56.024°N, 10.025°E, at a height of 142 m above sea level. The radar collects data up to a range of 240 km from the radar's location; however, the study catchment is located at a medium range of 36–103 km, which provides an ideal setup for high-quality radar measurements, avoiding the disadvantages of long-range and very short-range radar observations. A time series of 4 years of weather radar data was available.

The raw polar volume data were processed by initial filtering by the signal processor using signal quality index (SQI), Doppler and Laplacian of Gaussian speckle filters, and standard Doppler clutter filtering. The filtering is done to ensure optimal detection of precipitation for use in operational weather forecasting, which should also provide an equally high data quality for the hydrological application of this study.

The raw radar data were collected in polar volumes at a range bin size of 500 m, an angular beamwidth of 1° over nine elevation angles. Radar counts  $C$  recorded with 8-bit precision (values between 0 and 255) were converted to radar reflectivity  $Z$  (dBZ) using the following equation:

$$Z = 10^\gamma, \quad \text{where} \quad \gamma = \frac{1}{10}[\alpha(C - c_0) + \beta], \quad (1)$$

where  $\alpha$  is the gain,  $\beta$  is the intercept, and  $c_0$  is the offset ( $\alpha = 0.5$ ,  $\beta = -32$ , and  $c_0 = 0$ ). These values can be found in Table 1. Similarly, differential reflectivity  $Z_{\text{dr}}$  was converted from counts in the same way. Specific differential phase  $K_{\text{dp}}$  was not used as a parameter for

TABLE 1. Values of radar parameters used to convert from counts to dBZ.

	Gain	Offset	Value min	Value max
DBZH	0.5	−32	−32	95.5
ZDR	0.0625	−8	−8	7.94
PHIDP	1.417	−1.417	−1.417	359.9
RHOHV	0.005	0	0	1.27

estimating rainfall since a quality controlled  $K_{\text{dp}}$  product is still under development and evaluation.

Processing of radar data was done with the WRADLIB package, which is a Python-based open-source library that is freely available (Heistermann et al. 2013, 2015). A texture-based algorithm for ground clutter removal was deployed based on the method described in (Gabella and Notarpietro 2002), where nonstationary ground clutter and anomalously propagated echoes were identified and removed because they decorrelate with the surrounding pixels rapidly in space and time. Moreover, in the dual-pol databased rainfall product,  $\rho_{\text{hv}}$  was set with a threshold of 0.95 to remove the nonprecipitation echoes.

Attenuation corrections were applied to account for the signal power loss with range, where the process was done beam by beam, gate by gate using the algorithm proposed in Kramer and Verworn (2008). After clutter and attenuation corrections, radar data were interpolated from a polar grid to a 3D rectangular grid, and then constant altitude plan position indicators (CAPPIs) as well as pseudo-CAPPIs were calculated at 2 km height.

### b. Rainfall estimation algorithms

For the single-pol radar rainfall product, the traditional power-law Marshall–Palmer equation was used:

$$Z_h = AR^b, \quad R = (Z_h/A)^{1/b}, \quad (2)$$

where  $R$  is rainfall ( $\text{mm h}^{-1}$ ),  $Z_h$  is radar reflectivity in the horizontal direction, and  $A$  and  $b$  are empirical constants with values depending on the drop size distribution of the hydrometeors. In the present study, we used the values recommended by DMI, which are  $A = 220$  and  $b = 1.6$ , representing the most common rain type in Denmark.

For the dual-pol radar rainfall product, we chose to use an  $R(Z_h, Z_{\text{dr}})$  algorithm:

$$R = c_1 Z_h^{a_1} 10^{0.1b_1 Z_{\text{dr}}}, \quad (3)$$

where  $Z_{\text{dr}}$  is the radar differential reflectivity coming from the dual-pol radar. The terms  $a_1$ ,  $b_1$ , and  $c_1$  are coefficients that can be determined from regression analysis based on measured  $R$ ,  $Z_h$  and  $Z_{\text{dr}}$ . We have

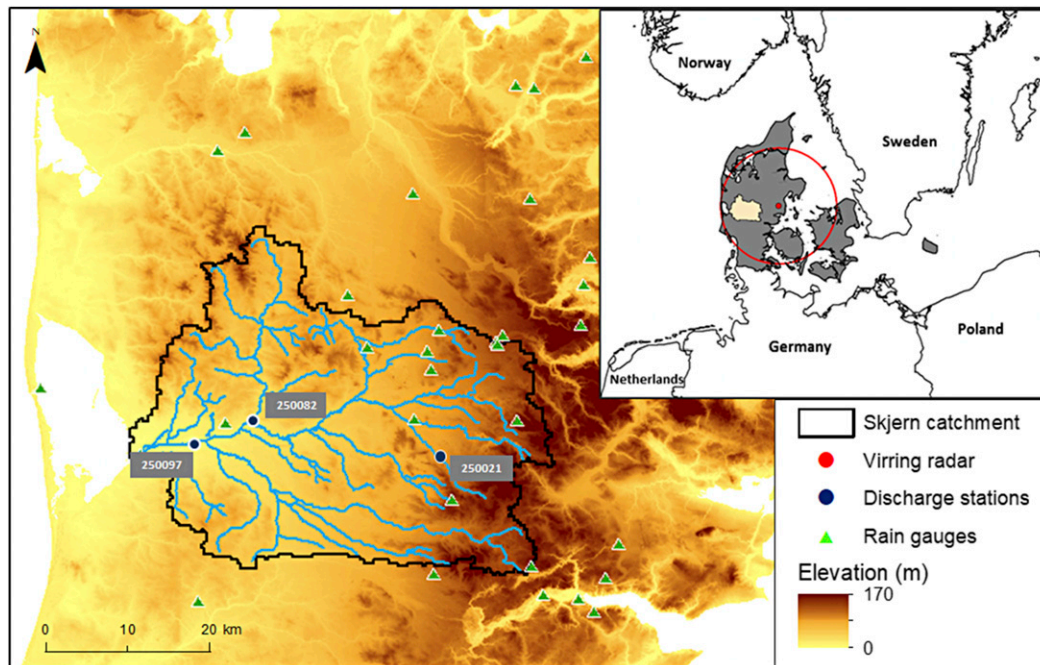


FIG. 1. Location of study area, Virring radar, rain gauges, and discharge stations.

chosen to use the coefficient values recommended for C-band radar in [Bringi and Chandrasekar \(2011\)](#), which are  $a_1 = 0.91$ ,  $b_1 = -2.09$ , and  $c_1 = 5.8 \times 10^{-3}$ .

For both the single-pol and the dual-pol radar rainfall estimation products, the hourly rainfall sum ( $\text{mm h}^{-1}$ ) was generated based on images from every 10 min. The hourly radar rainfall data were bias adjusted by hourly rain gauge data using the mean field bias (MFB) correction method:

$$\text{MFB} = \frac{\sum_{i=1}^n G_i}{\sum_{i=1}^n R_i}, \quad (4)$$

where  $G_i$  is the rain gauge observation, and  $R_i$  is the radar estimated value at the pixel that contains the rain gauge. Rain gauge data are collected from the database operated by the Danish hydrological observatory (HOBE; [Jensen and Illangasekare 2011](#)). There are in total 42 automatic gauges in the vicinity of the Skjern catchment as seen in [Fig. 1](#). Hourly rain gauge data are interpolated over the Skjern catchment using inverse distance weighting.

### c. Study area and hydrological model

The study features a modeling experiment of the Skjern catchment ([Fig. 1](#)), which is located in the western part of the Danish Jutland peninsula and covers around  $2500 \text{ km}^2$ . The various hydrological components of the catchment have been studied intensively at HOBE. The

climate of the catchment is characterized by maritime conditions with a mean annual precipitation of 990 mm and a mean reference evapotranspiration of 575 mm. The subsurface settings are predominately sandy and of glacial origin with intertwined sections of clay and till. The topography slopes gently from the coast to 125 m above sea level at the eastern boundary of the catchment. Agriculture is the predominant land use with around 70% areal coverage.

The hydrological model of the Skjern catchment is based on the MIKE Système Hydrologique Européen (MIKE SHE) code ([Abbott et al. 1986](#)). The modeling system consists of fully coupled modules of 3D groundwater flow based on Darcy's equation, 1D unsaturated flow based on Richards' equation, 1D river routing based on the kinematic wave approximation of St. Venant equations, and 2D overland flow based on the diffuse wave approximation of St. Venant equations. Moreover, the default MIKE SHE code is extended with an additional coupling of a two-component energy-balance-based land surface model, Shuttleworth and Wallace–Evapotranspiration (SW-ET; [Shuttleworth and Wallace 1985](#)). This extension enhances the physical representation of the processes taking place at the land–atmosphere interface where the imprint of precipitation variability is expected to be largest. The land surface model solves the energy balance at an hourly time step, and the diurnal variability can thus be described ([Overgaard 2005](#)). Besides hourly climate forcing data,



SW-ET requires a detailed vegetation parameterization (Stisen et al. 2011), which is derived from the Moderate Resolution Imaging Spectroradiometer (MODIS). Details are presented in Koch et al. (2017).

The Skjern model has a warm-up period from 2001 to 2010 and is subsequently run for a 4-yr simulation period starting in 2011. The model is run in an hourly time step. The horizontal discretization is 500 m. The model has previously undergone a parameter calibration against five independent observational datasets: stream discharge, hydraulic head, actual evapotranspiration, soil moisture, and satellite-derived LST. Details are presented in Stisen et al. (2018).

#### d. Pattern evaluation algorithm

The EOF analysis is a methodology commonly applied to evaluate large spatiotemporal datasets of soil moisture (Perry and Niemann 2007; Graf et al. 2014). The approach decomposes the variability of the dataset into two main components. First, a set of orthogonal spatial patterns (EOFs) are identified, which are time invariant and capture statistically significant patterns of covariation. Second, a set of loadings are computed that are time variant and specify the significance of each EOF over time. The mathematical background of the EOF methodology is described in more detail by Perry and Niemann (2008). Typically, the EOF analysis has been utilized to decompose the variability of a spatiotemporal dataset of a single hydrological variable with the aim to identify predominant modes of variability and their physical controls (Korres et al. 2010; Graf et al. 2014; Mascaro et al. 2015). To compare spatial patterns of two datasets, Koch et al. (2015) brought forward a novel concept of performing a joint EOF analysis on an integral data matrix that contains the two datasets to be compared. This extension of the traditional EOF analysis allows for a meaningful pattern similarity measure that was found to be insightful to quantify spatial model deficiencies (Koch et al. 2016b; Mendiguren et al. 2017). In this way, the resulting EOF maps honor the spatiotemporal variability of both datasets, and the weighted difference between the loadings at specific times can be utilized to derive a quantitative pattern similarity score.

Following this approach, the associated loadings of highly similar maps at a certain time step will show a minimal deviation. Vice versa, a large loading deviation can be attested to pattern dissimilarity. The loading deviation must be weighted according to the variance contribution of the corresponding EOF to ensure a reliable pattern skill score. Thus, the EOF-based similarity score  $S_{\text{EOF}}$  between an observed and a simulated map at time  $x$  can be formulated as

$$S_{\text{EOF}}^x = \sum_{i=1}^n w_i |(\text{load}_i^{\text{obsx}} - \text{load}_i^{\text{simx}})|, \quad (5)$$

where  $w_i$  is the variance contribution of the  $i$ th EOF,  $\text{load}^{\text{sim}}$  is the loading of the simulated pattern, and  $\text{load}^{\text{obs}}$  is loading of the observed pattern at time  $x$ . Prior to the EOF analysis, the mean is removed from each map, and thus the methodology is based on the spatial anomalies, which makes it bias insensitive.

### 3. Results

#### a. Estimated rainfall

An intense rainfall event is selected for demonstration purposes that spans 17–22 September 2012. This event is selected because it represents a typical rainfall in Denmark, namely, the stratiform rain lasting for a few days with occasional cloud burst. Four snapshots are taken at four different times, and the results are shown in Fig. 2. Since all gauges are reported wet at these times, the interpolated gauge product  $R(G)$  shows rainfall between all gauges. However, radar shows an entirely different picture for this event. There are in fact precipitating clouds covering some area of the catchment while leaving the rest areas dry. There are also clear spatial patterns indicating areas with intensive precipitation while the gauge-based product misses these spatial patterns completely. The two radar-based products,  $R(Z_h)$  based on a single-pol variable and  $R(Z_h, Z_{\text{dr}})$  based mainly on a dual-pol variable, exhibit only small noticeable differences, except that the high-intensity rainfall is more prevalent in  $R(Z_h, Z_{\text{dr}})$ . Such differences could be caused by the rainfall retrieval algorithm.

Figure 3 shows the comparison between estimated rainfall from the three products. The rainfall images are obtained on an hourly time step and are subsequently temporally aggregated to daily, monthly, and yearly values. Mean values of each image, either in its original form or aggregated, are calculated and plotted against the standard deviation of all pixels of that image. The figure provides insights into the spatiotemporal viability of rainfall. As expected, the spatial variability of rainfall is significantly decreased with temporal aggregation. At daily and subdaily time scales, the difference between rain-gauge- and radar-based precipitation is evident but not further noticeable at the monthly time scale and beyond. The radar products have an enhanced variability on the subdaily scale that is especially relevant for low-intensity rainfall events below  $0.1 \text{ mm day}^{-1}$ . When aggregated to yearly rainfall,  $R(G)$  exhibits the highest spatial variability, which can probably be caused by

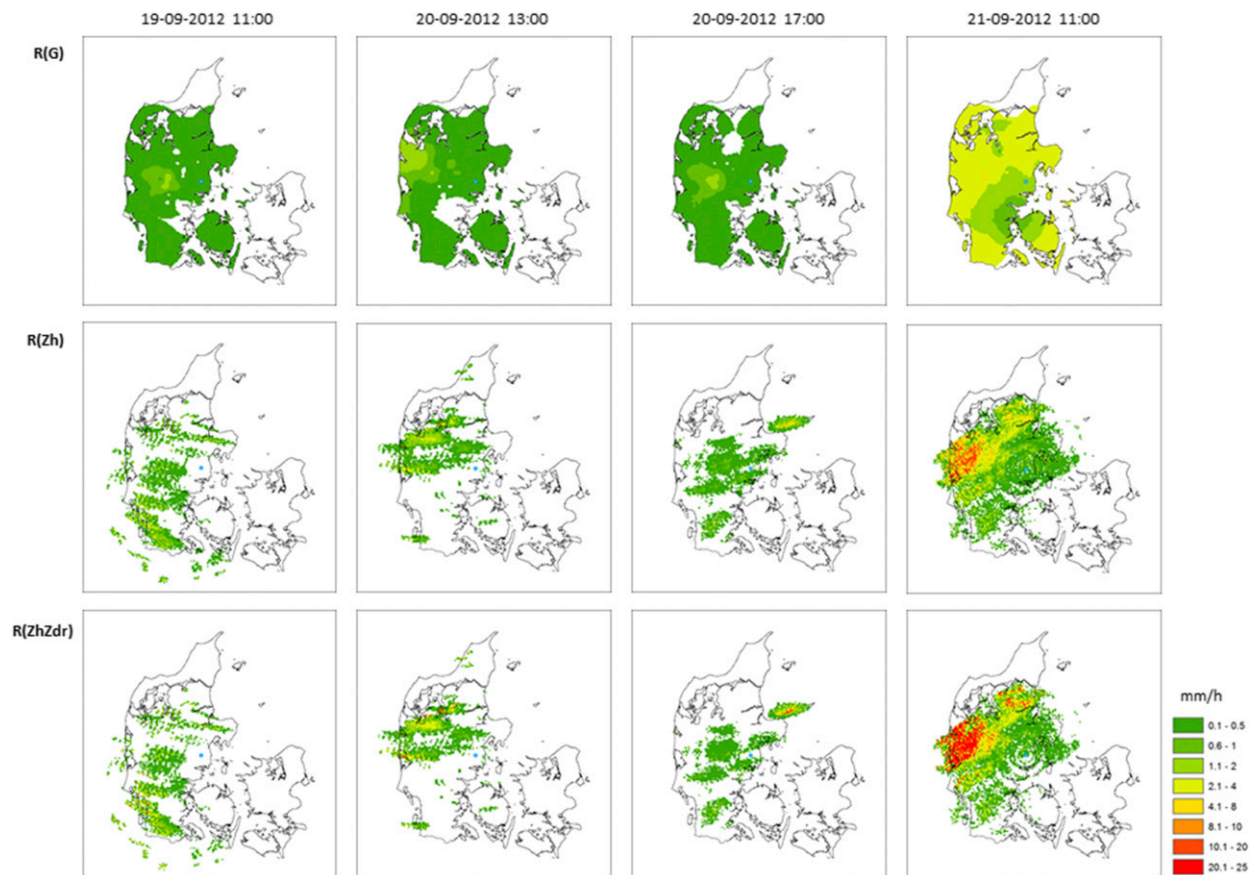


FIG. 2. Estimated precipitation for a rainfall event in September 2012. Columns show accumulated hourly precipitation at four times during the event. Rows show (top) quantitative precipitation estimation products using rain gauge data  $R(G)$ , (middle) single-pol radar data  $R(Z_h)$ , and (bottom) dual-pol radar data  $R(Z_h, Z_{dr})$ .

interpolation artifacts. There is no apparent difference between the two radar products. Parameter  $R(Z_h)$  has slightly higher spatial variability at daily and subdaily time scales most likely due to the removal of the nonprecipitating echoes in the  $R(Z_h, Z_{dr})$  product using the threshold value of  $\rho_{hv}$ . Therefore, implications for a hydrological simulation can be expected to take place at daily to subdaily time scales, whereas it can be expected that annual water budgets are not affected significantly by the rainfall products.

Probability density function (PDF) is an efficient way to better understand the rainfall intensity and frequency in the study area. PDF for both the hourly and daily rainfall data are produced, as seen in Fig. 4. We divided the dataset into 20 bins and plotted the rainfall intensity against the probability of occurrence in each bin. As shown in the figure, both the hourly and daily rainfall exhibits lognormal distributions. The peak of the daily rainfall PDF occurs at  $4 \text{ mm day}^{-1}$  with a probability of 16%. For hourly rainfall, the peak appears at  $0.22 \text{ mm h}^{-1}$ , with probability of 21%. The hourly rainfall has a sharper

peak and a longer tail than the daily rainfall PDF, which is expected due to the averaging effect of the daily mean values.

#### b. Simulated stream discharge and groundwater head

In Fig. 5 simulated stream discharge is plotted against observational data for two events, September 2011 and September 2012, and in Fig. 6 simulated groundwater heads averaged for the entire simulation period (2011–14) are shown. Moreover, model performance statistics are calculated using Nash–Sutcliffe efficiency (NSE) for stream discharge at station 250082 (shown in Fig. 1) and root-mean-square error (RMSE) for groundwater head in the main sandy aquifer over the entire catchment. Both statistics are done for the whole simulation period, and the results are seen in Table 2.

The table shows that the overall performance of all three models is acceptable, with scores comparable to previous studies in the same area (Stisen et al. 2018). It is suggested that at catchment scale and for long-term simulations, such as multiannual simulations, the spatial

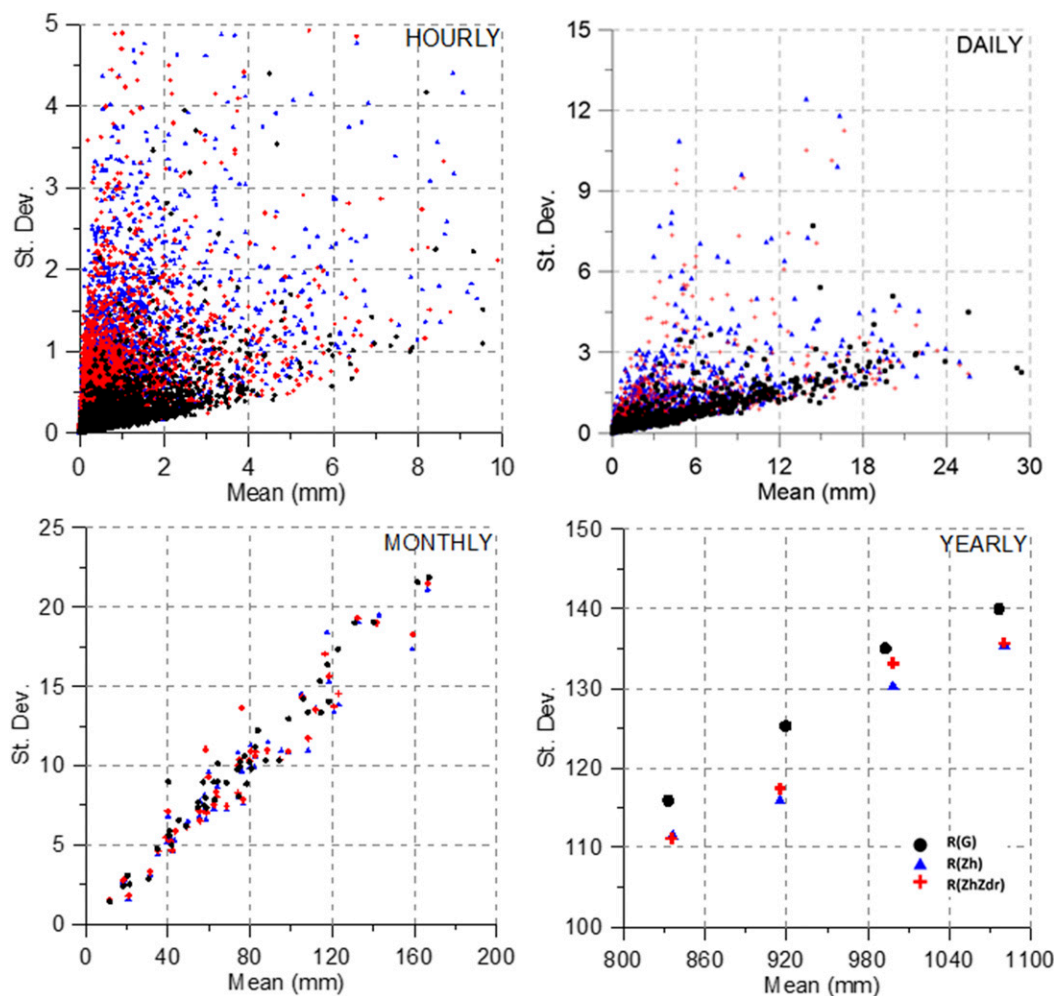


FIG. 3. Statistical analysis of estimated rainfall at hourly, daily, monthly, and yearly scale for the study period 2011–14 based on hourly rainfall data and aggregated to the various time periods. Mean values of each image are plotted against the standard deviation of all pixels of that image.

pattern of the precipitation forcing does not make a noticeable difference in the model results given that all precipitation products have been bias corrected in the same way. Likewise, the simulated groundwater heads for the three models are almost identical as seen in Fig. 6. Overall, the difference between the model evaluation scores is hardly noticeable for both surface water and groundwater.

The different precipitation forcings start to show their influence on the model simulations at subcatchment scale (station 250021), where the rain-gauge-based model exhibits different peak levels for the stream discharge flow, which could indicate that peak flow has higher uncertainty between models. Figure 5 also demonstrates the limitation of the observed data we used in this study for model evaluation because stream discharge is only observed at a daily time scale. Hence, the

benefits brought by running the model in hourly time step cannot be fully utilized, especially when the flow peaks occur in between of the observation points. The same problem happens to the groundwater data, where the typical observation frequency is days to months.

It is noted that the station with smallest catchment area has the worst model performance for both events, which is likely caused by how the model was calibrated. During model calibration, one important objective function was the so-called water balance at each stream gauging station. The goal was to reduce the bias to the minimum. As a result, much more weight is given to the streams with higher discharges or gauging stations that are located downstream, so that the overall water balance error can be smaller. As seen in Fig. 5, the highest peak volume for station 250021 is only 1/100 in comparison to the downstream station 250097. In addition,



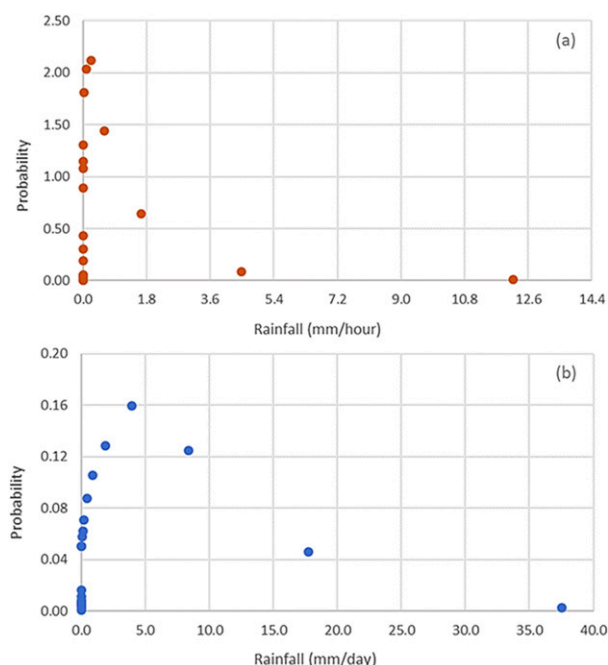


FIG. 4. PDFs calculated for rainfall at (a) hourly and (b) daily temporal scales using the rain gauge dataset for period 2011–14 in Skjern River catchment.

other studies also indicate that uncertainties are much larger at small scale than at large scale (He et al. 2011; Refsgaard et al. 2014). Oscillations are observed at the falling limb of the hydrographs in the September 2012 event. They are caused by numerical instabilities in the finite difference scheme used in the river routing module linked to the use of hourly rain data. The oscillations do not occur in the other event in September 2011.

### c. Pattern comparison

Figure 7 shows the calculated EOF scores using the single-pol radar-based model,  $R(Z_h)$ , as the benchmark. Again, the September event in 2012 is selected for demonstration purposes. An EOF score of zero indicates perfect agreement between two models, whereas higher EOF scores occur when the spatial patterns diverge due to differences in precipitation forcings. The selection of  $R(Z_h)$  as the benchmark is based on a simple logic: the calculation of EOF scores has to be the difference between two objects. If  $R(G)$  is chosen as the benchmark, then the two radar products cannot be compared with each other. Of the two radar products,  $R(Z_h)$  is preferred since it is expected to be the most commonly used one.

It is seen that during this rainfall event, pattern dissimilarity is highest at hours with high precipitation. The two radar products are similar in terms of general patterns despite the difference in intensity, whereas  $R(G)$

manifests a more distinct dissimilarity in spatial pattern. The simulated spatial patterns of LST and evapotranspiration (ET) are compared as well in order to characterize the effect of alternative precipitation products on spatial patterns of land surface variables. For ET,  $R(G)$  is again standing out as being more dissimilar in comparison to the two radar products. The pattern dissimilarity is most distinct in the first days where, despite low rainfall intensity, there seems to be significant differences in the precipitation patterns to cause diverging simulated patterns of ET. During the last day, the ET patterns are more or less identical, which is probably caused by the fact that it is raining across the entire catchment. In such a case the impact of differences between the rainfall products become indifferent because simulated ET reaches potential ET in most grids. LST, a variable closely linked to ET, shows a similar behavior. Recharge is highest at the last day of the rainfall event, leading to the most pronounced differences in spatial patterns between the models. During the first few days of the event, precipitation is not high enough to allow large amounts of water to propagate through the unsaturated zone to the saturated zone, which results in quite similar patterns of recharge between different models.

### d. Validation by remote sensing data

Remotely sensed LST data from the MODIS sensor on board the *Terra* and *Aqua* satellites are used to spatially evaluate the effect of diverging precipitation products for the applied catchment model (Fig. 8). The midday LST products from sensor MOD11A1 and MYD11A1 are used to obtain spatial patterns of LST. The LST product is at 1-km spatial resolution and acquisition time varies between 1100 and 1300 local time. We limit the spatial pattern evaluation to images that provide a cloud-free coverage of at least 90%. The analysis is based on 82 LST maps for the 4-yr simulation period. The EOF analysis indicates that there are no evident spatial differences between the different models. The remote sensing scenes are obtained under cloud-free conditions and hence not affected by rainfall variability, which leads to the distinctive spatial similarity between the simulated spatial patterns of LST.

### e. Spatial patterns and their scale dependency

Our results suggest that the spatial patterns of the various precipitation products and the impact on the hydrological responses simulated using a distributed hydrological model are highly scale dependent. To investigate the scaling problem in more detail, we have delineated the Skjern catchment into subcatchments with various sizes, as shown in Fig. 9, and the

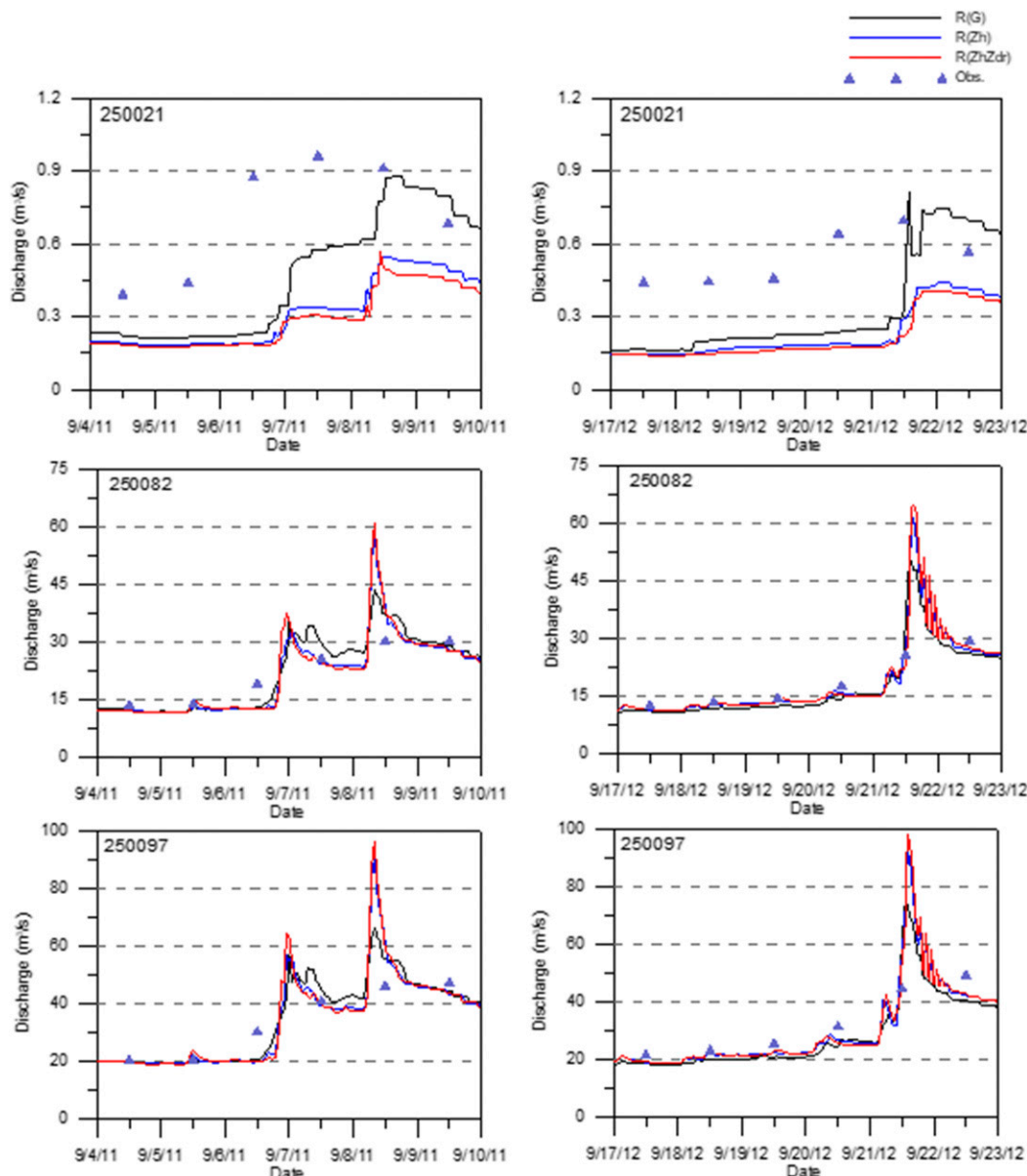


FIG. 5. Observed and simulated stream discharge for three discharge stations for two events: (left) September 2011 and (right) September 2012. Simulated discharge is hourly while observations are on daily basis.

corresponding EOF analysis is presented in Fig. 10. We use  $R(Z_h)$  as the benchmark and EOF scores are calculated for each individual subcatchment.

The mean EOF score in Fig. 10 relates to the mean of all hourly spatial pattern comparisons for the two rainfall events in September 2011 and September 2012. In general,  $R(G)$  simulates more dissimilar spatial patterns of hydrological response in contrast to  $R(Z_h, Z_{dr})$ . This is evident for the two variables, namely, ET and groundwater recharge. The variability is high for small subcatchments, which may relate to the fact that some

areas in the catchment are more affected by the spatial differences in the rainfall products, that is, the situation of having no rain in one product and rain in the other. Alternatively, this may also relate to the hydrological conditions of the subcatchments, where some may be located in the critical zone and thus have a strong coupling to the groundwater. Under such conditions, ET is controlled by the groundwater and variations in rainfall may affect the simulated spatial pattern of ET to a lesser degree. The EOF score seems to stabilize around  $100 \text{ km}^2$ , which indicates that rainfall variability is most

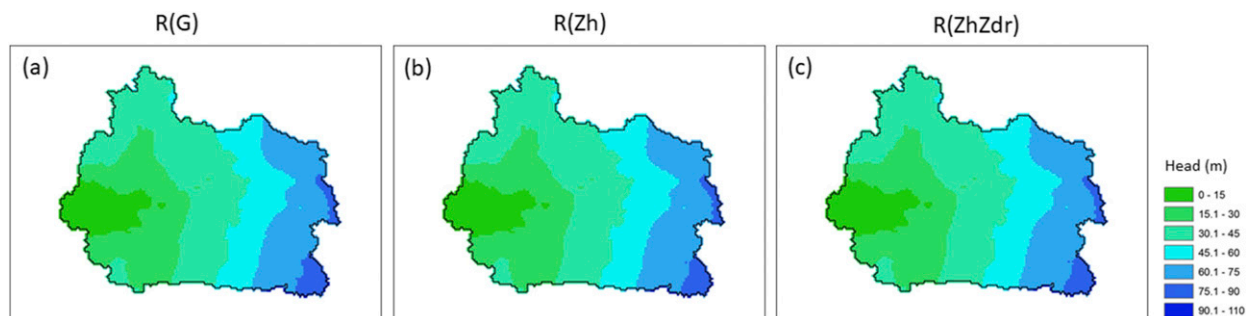


FIG. 6. Simulated average groundwater head (2011–14) based on the three rainfall products extracted from the fifth computational layer of the saturated zone representing the main sandy aquifer of the region.

crucial for subcatchments below this scale. Additionally, it needs to be pointed out that the results shown in the present research are more suitable for the wet season in temperate climate conditions with a flat terrain. They have not been tested for a different type of rainfall regime.

#### 4. Discussion and conclusions

It is our observation that the evaluation of distributed hydrological models has moved beyond an aggregated approach and toward more science-based spatially distributed methodologies (Stisen et al. 2011; Glaser et al. 2016). As a consequence, special emphasis should be put on the spatial dimension of model parameterization, forcing, calibration, and validation. The present study focuses on precipitation, the most dominant model forcing, and investigates differences between three products: 1) interpolated gauges  $R(G)$ , 2) single-pol radar  $R(Z_h)$ , and 3) dual-pol radar  $R(Z_h, Z_{dr})$ . Further, the hydrological implications of the different precipitation products on multiannual water budget simulations by an integrated catchment model are investigated in detail. Special attention is given to the comparisons of the spatial patterns in both the precipitation products and their corresponding hydrological responses. Our results confirm the previous notion that the traditional model evaluation criteria, such as using NSE and RMSE based on stream discharge and groundwater head observations, are not sufficient when natural patterns are highlighted in distributed hydrological modeling, since they are mostly insensitive to the simulated spatial patterns.

In our study area, namely, the Skjern River catchment, the terrain is rather flat. The highest point is about 125 m above sea level and sloping very gently from the hilltop to the sea. Therefore, it is not very common in the study area to experience orographic rain as in some mountainous regions in others parts of Europe.

If complex terrain had existed, first of all, we would expect that the rainfall systems would have enhanced spatial variability due to the orographic effects, and thus enhanced local hydrological response in terms of increased total flood volume (Buytaert et al. 2006; Delrieu et al. 2009). In such cases, the rain gauge density in the Skjern catchment would not reach the minimum requirement to capture the spatial variability of rainfall, and radar-based QPE products would play a more significant role in storm simulations. Second, complex terrain often poses challenges to the radar-based QPE due to partial or total radar beam blockage. Radar variables that are measurements of the magnitude of the return echoes, such as  $Z_h$  and  $Z_{dr}$ , would be significantly affected. Third, some studies have indicated that the specific differential phase  $K_{dp}$  was immune to beam blockage (Zrnić and Ryzhkov 1996; Friedrich et al. 2007). However, others have suggested that  $K_{dp}$  data should be corrected for the nonuniform vertical profile before being used in the rainfall estimation (Wang et al. 2013). In any case, we would anticipate an increase of radar QPE accuracy by including more dual-pol variables when complex terrain is present.

Satellite data are used to obtain LST images so that we have an independent data source to verify the simulated spatial patterns of hydrological responses driven by radar- and rain-gauge-based rainfall estimates. Although not used in the present study, satellite-based rainfall retrieval is a very commonly used approach when it comes to spatial rainfall estimation. It is acknowledged that a rain gauge is the only direct

TABLE 2. Hydrological model performance evaluation of all the three models using NSE for stream discharge and RMSE for groundwater head respectively.

	$R(G)$	$R(Z_h)$	$R(Z_h, Z_{dr})$
NSE	0.882	0.894	0.900
RMSE	3.301	3.322	3.289

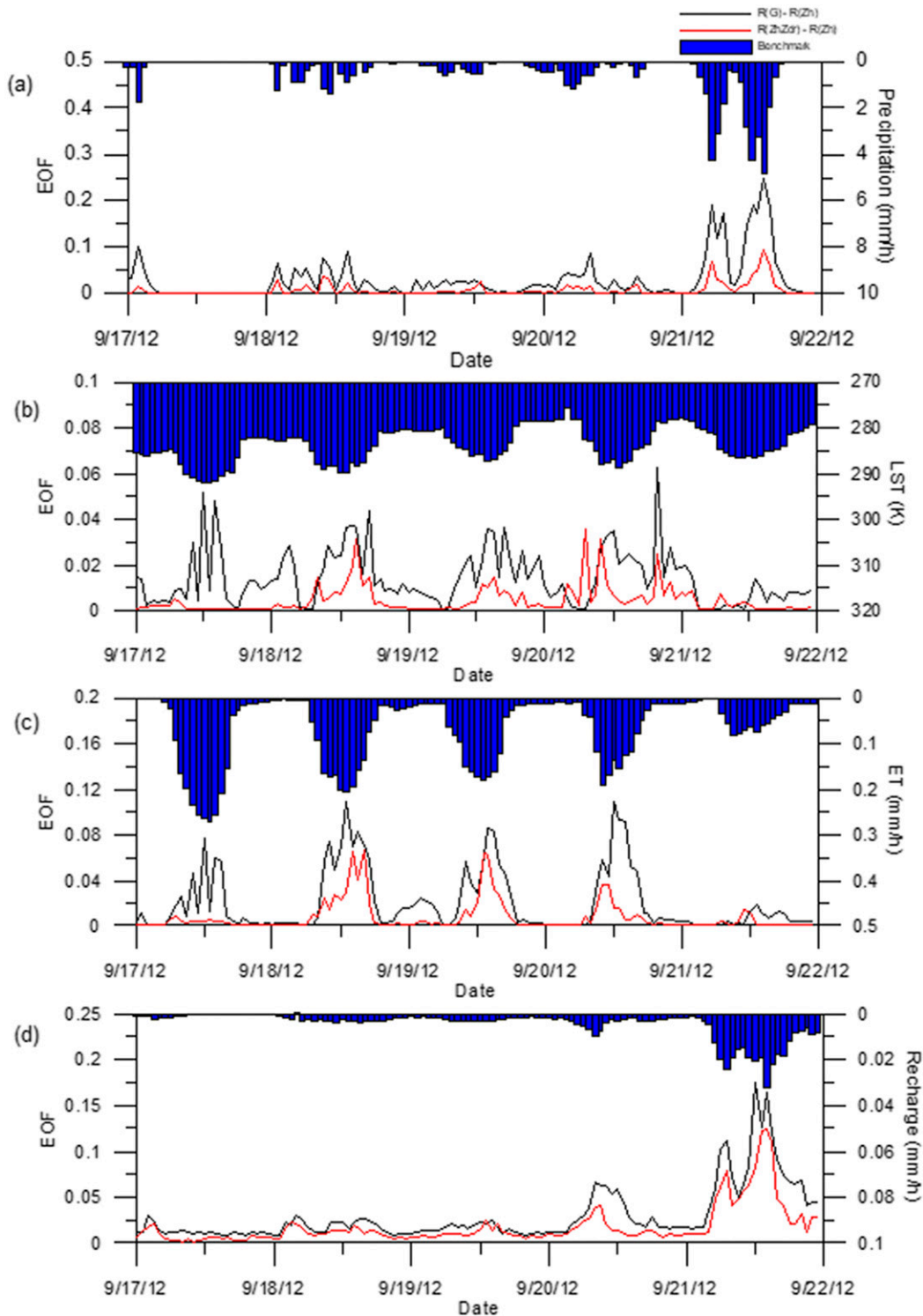


FIG. 7. EOF analysis of hydrological forcing and responses for the September 2012 rainfall event: (a) precipitation, (b) LST, (c) ET, and (d) groundwater recharge. The  $R(Z_h)$ -based model is used as benchmark, and the EOF scores reflect the pattern similarity between  $R(Z_h)$  and the corresponding model.

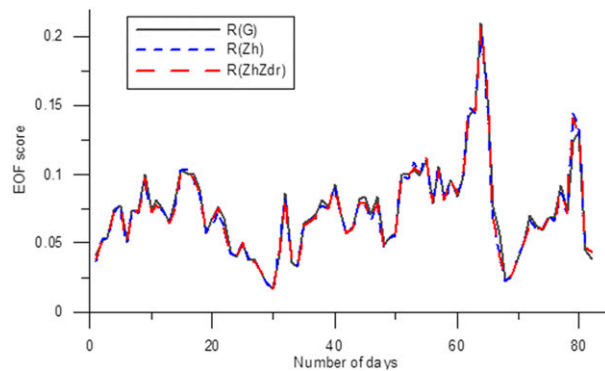


FIG. 8. Validation of simulated LST patterns for different precipitation products against LST patterns based on remotely sensed MODIS images.

measurement of rainfall, while both radar and satellite are remote sensing instruments that only give indirect approximations of rainfall. Rain gauges are usually too sparsely placed in order to capture the complete rainfall spatial variability in most places of the world, but their measurements at point scale are relatively trustworthy. Comparing radar and satellite, radar usually has higher spatial and temporal resolutions, and it is closer to the

elevation where precipitation actually takes place. However, radar also suffers from various sources of uncertainties such as signal attenuation, false echoes, complex terrain effect, etc. Radars are also rather costly and not available in some of the developing nations. For the above reasons, one could argue that radar data are more suitable for filling gaps in between the gauges while satellite data are more suitable for filling in between radar estimates. Several studies use either rain-gauge- or radar-based rainfall estimation as the “ground level truth” to benchmark satellite rainfall (Anagnostou et al. 2010; Cimini et al. 2013). There are other studies that use all available sources of information, which include rain gauge, radar, and satellite, to create data fusion products. These data fusion products have shown their functionality in estimating large-scale precipitation systems in longer time series (He et al. 2018).

With the employment of the EOF analysis, we are able to identify pattern similarities/dissimilarities, and thus the hydrological implications. By comparing different precipitation products, including radar- and rain-gauge-based products and different radar algorithms, it is found that the difference between their spatial patterns is related to rainfall signal intensity, as well as the scales in time and

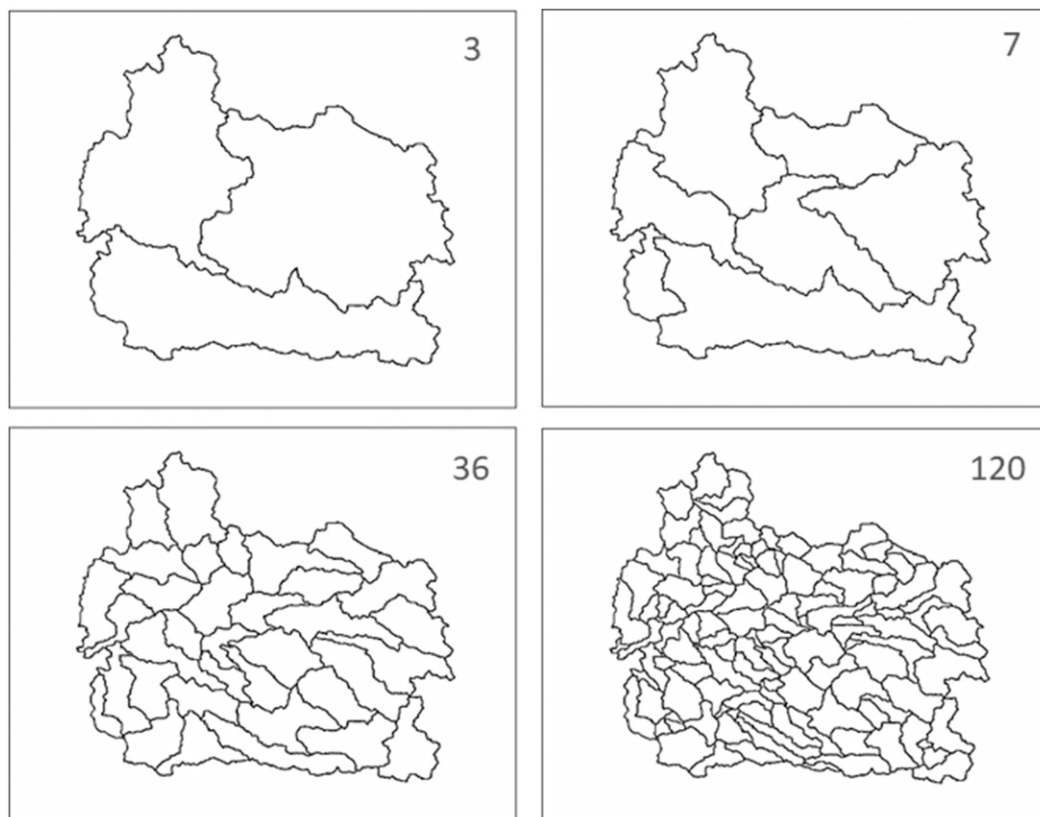


FIG. 9. Delineation of 166 subcatchments within the Skjern catchment. Subcatchment areas range from 3 to 1030 km<sup>2</sup>.



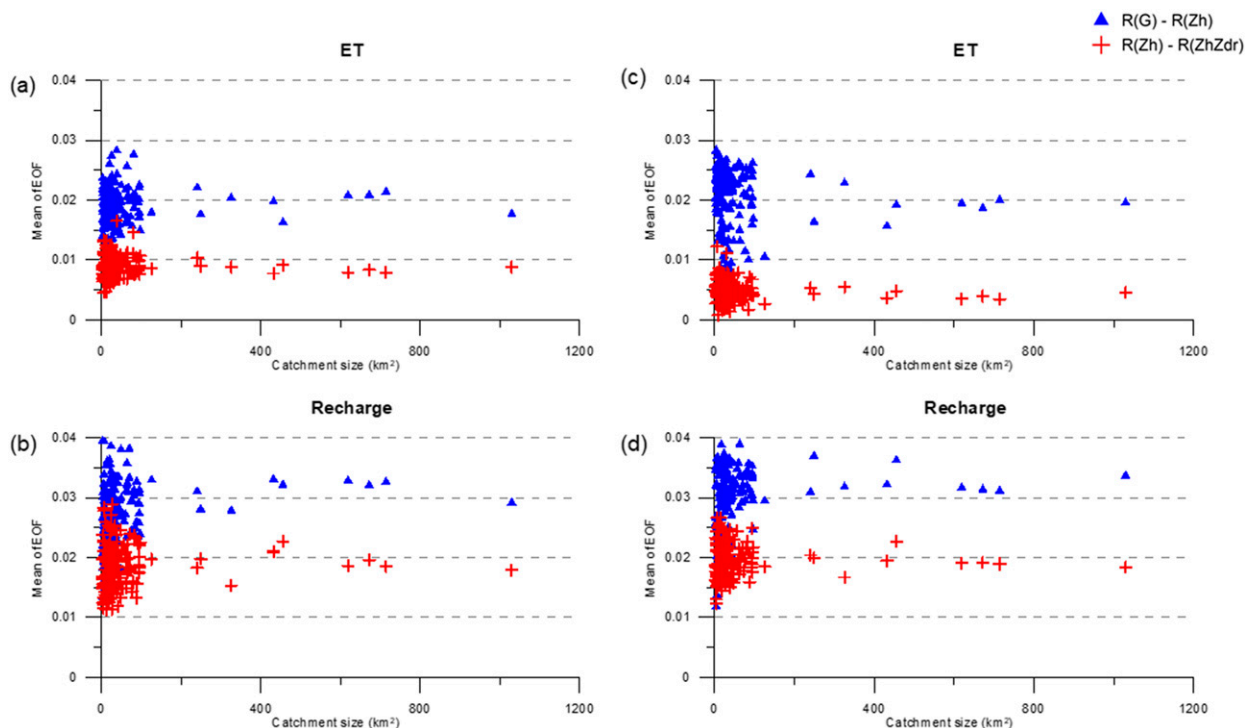


FIG. 10. Spatial-scale analysis of simulated hydrological responses for the (a),(b) September 2011 event and (c),(d) September 2012 event based on 166 subcatchments and using the  $R(Z_h)$  model as benchmark.

space. To be more specific, higher precipitation intensity usually leads to higher pattern dissimilarity, since it makes a big difference by switching the high-intensity cells just a few grids. In addition, the intensity of hydrological response is naturally linked to the rainfall intensity that makes such events more prone to pattern dissimilarity. For the temporal-scale dependency, the pattern differences are noticeable at hourly and daily scales where the radar- and rain-gauge-based precipitation maps can be entirely different. However, when it comes to monthly and yearly time scales, the differences start to disappear. For the spatial-scale dependency, the simulated spatial patterns of hydrological response are noticeable below  $100 \text{ km}^2$ , which is consistent with previous findings (He et al. 2011).

Our results suggest that different components of the hydrological cycle respond to the diverging rainfall forcing patterns at different temporal lags. Land surface variables such as ET or LST respond quickly to single rainfall events, but then the reaction becomes slower as the soil gets saturated. Groundwater recharge, on the other hand, remains unaffected during the start of the rain event when precipitation is still small. The impact first emerges toward the end of the event and requires also stronger rainfall intensity.

In the present study, the EOF analysis is applied three times: first, to compare the hourly precipitation patterns of

the three products during an intense rainfall event; second, to evaluate the spatial performance of the three competing models against observed LST; and third, to perform a scaling analysis to demonstrate the model spatial similarity at different subcatchment scales. We regard the EOF approach as a reliable and insightful metric to quantify the spatial similarity of patterns of hydrological variables. Its bias insensitivity is considered favorable, because it focuses on the internal distribution of catchment response and not on magnitude. There are comparable methods also available for spatial pattern analysis, such as fractions skill score and the like (Roberts and Lean 2008; Gilleland et al. 2009). The aim of the current study is not to evaluate the reliability of the EOF methodology, but to use it as a tool to compare alternative spatial patterns.

Between the single-pol and dual-pol radar data, successful stories have been reported when dual-pol data outperform both meteorologically and hydrologically, when comparing to observed rainfall data or observed streamflow data (Cunha et al. 2013; Seo et al. 2015; Chen et al. 2017). However, our study experiences difficulties in reaching the same conclusion, perhaps because of the lack of data. Rainfall spatial patterns exhibit visible differences; however, the validation by using the remote sensing data is somewhat indecisive since the observed LST data are not available at subdaily time steps and are

limited to cloud-free conditions. Groundwater head data are distributed but temporally sparse. Therefore, in order to demonstrate the difference between single-pol and dual-pol data in an integrated catchment modeling context and subsequently validate simulated hydrological patterns, finer-resolution data are urgently needed.

The goal of the present study is to focus on the difference between single-pol- and dual-pol-based radar rainfall estimation from both spatial and temporal scales as well as the pattern similarities. There are numerous radar rainfall retrieval algorithms available (e.g., Kalogiros et al. 2013; Ryzhkov et al. 2014). However, the selection of rainfall retrieval algorithm is not given too much attention.

Knowing that when it comes to estimation of rainfall both mean and spatial pattern are equally important, we single out the spatial pattern issues where the rainfall mean has been intentionally neglected by using the MFB correction. Whether the bias corrected rain gauge data are able to provide the true mean values is outside the scope of the present study. It has been advocated by several researchers that QPE should be carried out without any help from the rain gauge data in a hydrological context (Cifelli et al. 2011; Chang et al. 2016). However, such idea is still debatable, and if MFB correction is not performed, our results could be significantly changed.

The present study demonstrates the challenges and potentials when balancing the complex, spatially distributed nature of our models and the available data both for model forcing and validation. It is, to our knowledge, the first comparison of single-pol and dual-pol radar rainfall estimation and their hydrological responses with a focus on spatial patterns. However, several issues have not been covered that are worth pursuing for further analysis in future studies, such the radar rainfall algorithms and/or the hydrological model code selected.

**Acknowledgments.** This work has been supported by the National Key R&D Program of China, Project 2016YFC0401404; the National Natural Science Foundation of China (NSFC), Grant 41701042; and a grant from the VILLUM Foundation to the Danish Hydrological Observatory HOBE. The authors thank Professor Jens Christian Resgaard from the Geological Survey of Denmark and Greenland for his kind help during the writing of this paper. The authors would like to thank the Shenzhen Municipal Engineering Lab of Environmental IoT Technologies for the support.

## REFERENCES

- Abbott, M. B., J. C. Bathurst, J. A. Cunge, P. E. O'Connell, and J. Rasmussen, 1986: An introduction to the European Hydrological System—Système Hydrologique Européen, “SHE”, 2: Structure of a physically-based, distributed modelling system. *J. Hydrol.*, **87**, 61–77, [https://doi.org/10.1016/0022-1694\(86\)90115-0](https://doi.org/10.1016/0022-1694(86)90115-0).
- Anagnostou, E. N., V. Maggioni, E. I. Nikolopoulos, T. Meskele, F. Hossain, and A. Papadopoulos, 2010: Benchmarking high-resolution global satellite rainfall products to radar and rain-gauge rainfall estimates. *IEEE Trans. Geosci. Remote Sens.*, **48**, 1667–1683, <https://doi.org/10.1109/TGRS.2009.2034736>.
- Arsenault, R., and F. Brissette, 2014: Determining the optimal spatial distribution of weather station networks for hydrological modeling purposes using RCM datasets: An experimental approach. *J. Hydrometeor.*, **15**, 517–526, <https://doi.org/10.1175/JHM-D-13-088.1>.
- Bárdossy, A., and T. Das, 2008: Influence of rainfall observation network on model calibration and application. *Hydrol. Earth Syst. Sci.*, **12**, 77–89, <https://doi.org/10.5194/hess-12-77-2008>.
- Berne, A., and W. F. Krajewski, 2013: Radar for hydrology: Unfulfilled promise or unrecognized potential? *Adv. Water Resour.*, **51**, 357–366, <https://doi.org/10.1016/j.advwatres.2012.05.005>.
- Boyle, D. P., H. V. Gupta, S. Sorooshian, V. Koren, Z. Y. Zhang, and M. Smith, 2001: Toward improved streamflow forecasts: Value of semidistributed modeling. *Water Resour. Res.*, **37**, 2749–2759, <https://doi.org/10.1029/2000WR000207>.
- Bringi, V. N., and V. Chandrasekar, 2011: *Polarimetric Doppler Weather Radar: Principles and Applications*. Cambridge University Press, 664 pp.
- Buytaert, W., R. Cellier, P. Willems, B. De Bievre, and G. Wyseure, 2006: Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *J. Hydrol.*, **329**, 413–421, <https://doi.org/10.1016/j.jhydrol.2006.02.031>.
- Carpenter, T. M., J. A. Sperflage, K. P. Georgakakos, T. Sweeney, and D. L. Fread, 1999: National threshold runoff estimation utilizing GIS in support of operational flash flood warning systems. *J. Hydrol.*, **224**, 21–44, [https://doi.org/10.1016/S0022-1694\(99\)00115-8](https://doi.org/10.1016/S0022-1694(99)00115-8).
- Chang, W. Y., J. Vivekanandan, K. Ikeda, and P. L. Lin, 2016: Quantitative precipitation estimation of the epic 2013 Colorado flood event: Polarization radar-based variational scheme. *J. Appl. Meteor. Climatol.*, **55**, 1477–1495, <https://doi.org/10.1175/JAMC-D-15-0222.1>.
- Chen, H. N., V. Chandrasekar, and R. Bechini, 2017: An improved dual-polarization radar rainfall algorithm (DROPS2.0): Application in NASA IFloodS field campaign. *J. Hydrometeor.*, **18**, 917–937, <https://doi.org/10.1175/JHM-D-16-0124.1>.
- Cifelli, R., V. Chandrasekar, S. Lim, P. C. Kennedy, Y. Wang, and S. A. Rutledge, 2011: A new dual-polarization radar rainfall algorithm: Application in Colorado precipitation events. *J. Atmos. Oceanic Technol.*, **28**, 352–364, <https://doi.org/10.1175/2010JTECHA1488.1>.
- Cimini, D., and Coauthors, 2013: Validation of satellite OPEMW precipitation product with ground-based weather radar and rain gauge networks. *Atmos. Meas. Tech.*, **6**, 3181–3196, <https://doi.org/10.5194/amt-6-3181-2013>.
- Clark, M. P., D. Kavetski, and F. Fenicia, 2011: Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resour. Res.*, **47**, W09301, <https://doi.org/10.1029/2010WR009827>.
- Cole, S. J., and R. J. Moore, 2008: Hydrological modelling using raingauge- and radar-based estimators of areal rainfall. *J. Hydrol.*, **358**, 159–181, <https://doi.org/10.1016/j.jhydrol.2008.05.025>.
- , and —, 2009: Distributed hydrological modelling using weather radar in gauged and ungauged basins. *Adv. Water Resour.*, **32**, 1107–1120, <https://doi.org/10.1016/j.advwatres.2009.01.006>.

- Cunha, L. K., J. A. Smith, M. L. Baeck, and W. F. Krajewski, 2013: An early performance evaluation of the NEXRAD dual-polarization radar rainfall estimates for urban flood applications. *Wea. Forecasting*, **28**, 1478–1497, <https://doi.org/10.1175/WAF-D-13-00046.1>.
- Delrieu, G., B. Boudevillain, J. Nicol, B. Chapon, P. E. Kirstetter, H. Andrieu, and D. Faure, 2009: Bollène-2002 Experiment: Radar quantitative precipitation estimation in the Cévennes-Vivarais region, France. *J. Appl. Meteor. Climatol.*, **48**, 1422–1447, <https://doi.org/10.1175/2008JAMC1987.1>.
- Dong, X. H., C. M. Dohmen-Janssen, and M. J. Booij, 2005: Appropriate spatial sampling of rainfall for flow simulation. *Hydrol. Sci. J.*, **50**, 279–298, <https://doi.org/10.1623/hysj.50.2.279.61801>.
- Fang, Z. F., H. Bogen, S. Kollet, J. Koch, and H. Vereecken, 2015: Spatio-temporal validation of long-term 3D hydrological simulations of a forested catchment using empirical orthogonal functions and wavelet coherence analysis. *J. Hydrol.*, **529**, 1754–1767, <https://doi.org/10.1016/j.jhydrol.2015.08.011>.
- Friedrich, K., U. Germann, J. J. Gourley, and P. Tabary, 2007: Effects of radar beam shielding on rainfall estimation for the polarimetric C-band radar. *J. Atmos. Oceanic Technol.*, **24**, 1839–1859, <https://doi.org/10.1175/JTECH2085.1>.
- Gabella, M., and R. Notarpietro, 2002: Ground clutter characterization and elimination in mountainous terrain. *Proc. Second European Conf. on Radar in Meteorology and Hydrology*, Delft, Netherlands, Delft University of Technology, 305–311, <https://www.copernicus.org/erad/online/erad-305.pdf>.
- Gao, Y. F., Y. D. Chen, L. N. Zhang, and T. Peng, 2016: Radar-rainfall estimation from S-band radar and its impact on the runoff simulation of a heavy rainfall event in the Huaihe River basin. *J. Meteor. Soc. Japan*, **94**, 75–89, <https://doi.org/10.2151/jmsj.2016-005>.
- Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of spatial forecast verification methods. *Wea. Forecasting*, **24**, 1416–1430, <https://doi.org/10.1175/2009WAF2222269.1>.
- Glaser, B., J. Klaus, S. Frei, J. Frentress, L. Pfister, and L. Hopp, 2016: On the value of surface saturated area dynamics mapped with thermal infrared imagery for modeling the hillslope-river-stream continuum. *Water Resour. Res.*, **52**, 8317–8342, <https://doi.org/10.1002/2015WR018414>.
- Goudenhoofd, E., and L. Delobbe, 2016: Generation and verification of rainfall estimates from 10-yr volumetric weather radar measurements. *J. Hydrometeorol.*, **17**, 1223–1242, <https://doi.org/10.1175/JHM-D-15-0166.1>.
- Gourley, J. J., S. E. Giangrande, Y. Hong, Z. L. Flamig, T. Schuur, and J. A. Vrugt, 2010: Impacts of polarimetric radar observations on hydrologic simulation. *J. Hydrometeorol.*, **11**, 781–796, <https://doi.org/10.1175/2010JHM1218.1>.
- Graf, A., H. R. Bogen, C. Drue, H. Hardelauf, T. Putz, G. Heinemann, and H. Vereecken, 2014: Spatiotemporal relations between water budget components and soil water content in a forested tributary catchment. *Water Resour. Res.*, **50**, 4837–4857, <https://doi.org/10.1002/2013WR014516>.
- Grayson, R. B., G. Bloschl, A. W. Western, and T. A. McMahon, 2002: Advances in the use of observed spatial patterns of catchment hydrological response. *Adv. Water Resour.*, **25**, 1313–1334, [https://doi.org/10.1016/S0309-1708\(02\)00060-X](https://doi.org/10.1016/S0309-1708(02)00060-X).
- He, X., J. C. Refsgaard, T. O. Sonnenborg, F. Vejen, and K. H. Jensen, 2011: Statistical analysis of the impact of radar rainfall uncertainties on water resources modeling. *Water Resour. Res.*, **47**, W09526, <https://doi.org/10.1029/2011WR010670>.
- , T. O. Sonnenborg, J. C. Refsgaard, F. Vejen, and K. H. Jensen, 2013: Evaluation of the value of radar QPE data and rain gauge data for hydrological modeling. *Water Resour. Res.*, **49**, 5989–6005, <https://doi.org/10.1002/wrcr.20471>.
- He, Y. X., Y. Zhang, R. Kuligowski, R. Cifelli, and D. Kitzmiller, 2018: Incorporating satellite precipitation estimates into a radar-gauge multi-sensor precipitation estimation algorithm. *Remote Sens.*, **10**, 106, <https://doi.org/10.3390/rs10010106>.
- Heistermann, M., S. Jacobi, and T. Pfaff, 2013: Technical note: An open source library for processing weather radar data (*wradlib*). *Hydrol. Earth Syst. Sci.*, **17**, 863–871, <https://doi.org/10.5194/hess-17-863-2013>.
- , and Coauthors, 2015: The emergence of open-source software for the weather radar community. *Bull. Amer. Meteor. Soc.*, **96**, 117–128, <https://doi.org/10.1175/BAMS-D-13-00240.1>.
- Immerzeel, W. W., and P. Droogers, 2008: Calibration of a distributed hydrological model based on satellite evapotranspiration. *J. Hydrol.*, **349**, 411–424, <https://doi.org/10.1016/j.jhydrol.2007.11.017>.
- Jensen, K. H., and T. H. Illangasekare, 2011: Hobe: A hydrological observatory. *Vadose Zone J.*, **10**, 1–7, <https://doi.org/10.2136/vzj2011.0006>.
- Kalogiros, J., M. N. Anagnostou, E. N. Anagnostou, M. Montopoli, E. Picciotti, and F. S. Marzano, 2013: Optimum estimation of rain microphysical parameters from X-band dual-polarization radar observables. *IEEE Trans. Geosci. Remote Sens.*, **51**, 3063–3076, <https://doi.org/10.1109/TGRS.2012.2211606>.
- Koch, J., K. H. Jensen, and S. Stisen, 2015: Toward a true spatial model evaluation in distributed hydrological modeling: Kappa statistics, Fuzzy theory, and EOF-analysis benchmarked by the human perception and evaluated against a modeling case study. *Water Resour. Res.*, **51**, 1225–1246, <https://doi.org/10.1002/2014WR016607>.
- , T. Cornelissen, Z. F. Fang, H. Bogen, B. Diekkruiger, S. Kollet, and S. Stisen, 2016a: Inter-comparison of three distributed hydrological models with respect to seasonal variability of soil moisture patterns at a small forested catchment. *J. Hydrol.*, **533**, 234–249, <https://doi.org/10.1016/j.jhydrol.2015.12.002>.
- , A. Siemann, S. Stisen, and J. Sheffield, 2016b: Spatial validation of large-scale land surface models against monthly land surface temperature patterns using innovative performance metrics. *J. Geophys. Res. Atmos.*, **121**, 5430–5452, <https://doi.org/10.1002/2015JD024482>.
- , G. Mendiguren, G. Mariethoz, and S. Stisen, 2017: Spatial sensitivity analysis of simulated land surface patterns in a catchment model using a set of innovative spatial performance metrics. *J. Hydrometeorol.*, **18**, 1121–1142, <https://doi.org/10.1175/JHM-D-16-0148.1>.
- Korres, W., C. N. Koyama, P. Fiener, and K. Schneider, 2010: Analysis of surface soil moisture patterns in agricultural landscapes using empirical orthogonal functions. *Hydrol. Earth Syst. Sci.*, **14**, 751–764, <https://doi.org/10.5194/hess-14-751-2010>.
- Kramer, S., and H.-R. Verworn, 2008: Improved C-band radar data processing for real time control of urban drainage systems. *Eleventh Int. Conf. on Urban Drainage*, Edinburgh, United Kingdom, IAHR/IWA, 10 pp., [https://web.sbe.hw.ac.uk/staffprofiles/bdgsa/11th\\_International\\_Conference\\_on\\_Urban\\_Drainage\\_CD/ICUD08/pdfs/105.pdf](https://web.sbe.hw.ac.uk/staffprofiles/bdgsa/11th_International_Conference_on_Urban_Drainage_CD/ICUD08/pdfs/105.pdf).
- Kuczera, G., and B. J. Williams, 1992: Effect of rainfall errors on accuracy of design flood estimates. *Water Resour. Res.*, **28**, 1145–1153, <https://doi.org/10.1029/91WR03002>.
- Li, H. T., P. Brunner, W. Kinzelbach, W. P. Li, and X. G. Dong, 2009: Calibration of a groundwater model using pattern

- information from remote sensing data. *J. Hydrol.*, **377**, 120–130, <https://doi.org/10.1016/j.jhydrol.2009.08.012>.
- Mascaro, G., E. R. Vivoni, and L. A. Mendez-Barroso, 2015: Hyperresolution hydrologic modeling in a regional watershed and its interpretation using empirical orthogonal functions. *Adv. Water Resour.*, **83**, 190–206, <https://doi.org/10.1016/j.advwatres.2015.05.023>.
- Mausson, F., D. Scherer, T. Molg, E. Collier, J. Curio, and R. Finkelburg, 2014: Precipitation seasonality and variability over the Tibetan Plateau as resolved by the High Asia Reanalysis. *J. Climate*, **27**, 1910–1927, <https://doi.org/10.1175/JCLI-D-13-00282.1>.
- Mendiguren, G., J. Koch, and S. Stisen, 2017: Spatial pattern evaluation of a calibrated national hydrological model—A remote-sensing-based diagnostic approach. *Hydrol. Earth Syst. Sci.*, **21**, 5987–6005, <https://doi.org/10.5194/hess-21-5987-2017>.
- Orth, R., E. Dutra, I. F. Trigo, and G. Balsamo, 2017: Advancing land surface model development with satellite-based earth observations. *Hydrol. Earth Syst. Sci.*, **21**, 2483–2495, <https://doi.org/10.5194/hess-21-2483-2017>.
- Overgaard, J., 2005: Energy-based land-surface modelling: New opportunities in integrated hydrological modelling. Ph.D. thesis, Institute of Environment and Resources, Technical University of Denmark, 39 pp.
- Perry, M. A., and J. D. Niemann, 2007: Analysis and estimation of soil moisture at the catchment scale using EOFs. *J. Hydrol.*, **334**, 388–404, <https://doi.org/10.1016/j.jhydrol.2006.10.014>.
- , and —, 2008: Generation of soil moisture patterns at the catchment scale by EOF interpolation. *Hydrol. Earth Syst. Sci.*, **12**, 39–53, <https://doi.org/10.5194/hess-12-39-2008>.
- Refsgaard, J. C., and Coauthors, 2014: Nitrate reduction in geologically heterogeneous catchments—A framework for assessing the scale of predictive capability of hydrological models. *Sci. Total Environ.*, **468–469**, 1278–1288, <https://doi.org/10.1016/j.scitotenv.2013.07.042>.
- Renard, P., and D. Allard, 2013: Connectivity metrics for subsurface flow and transport. *Adv. Water Resour.*, **51**, 168–196, <https://doi.org/10.1016/j.advwatres.2011.12.001>.
- Rice, J. S., R. E. Emanuel, J. M. Vose, and S. A. C. Nelson, 2015: Continental U.S. streamflow trends from 1940 to 2009 and their relationships with watershed spatial characteristics. *Water Resour. Res.*, **51**, 6262–6275, <https://doi.org/10.1002/2014WR016367>.
- Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, <https://doi.org/10.1175/2007MWR2123.1>.
- Ruiz-Pérez, G., J. Koch, S. Manfreda, K. Caylor, and F. Francés, 2017: Calibration of a parsimonious distributed ecohydrological daily model in a data-scarce basin using exclusively the spatio-temporal variation of NDVI. *Hydrol. Earth Syst. Sci.*, **21**, 6235–6251, <https://doi.org/10.5194/hess-21-6235-2017>.
- Ryzhkov, A., D. Zrníc, and D. Atlas, 1997: Polarimetrically tuned  $R(Z)$  relations and comparison of radar rainfall methods. *J. Appl. Meteor.*, **36**, 340–349, [https://doi.org/10.1175/1520-0450\(1997\)036<0340:PTRZRA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1997)036<0340:PTRZRA>2.0.CO;2).
- , T. J. Schuur, D. W. Burgess, P. L. Heinselman, S. E. Giangrande, and D. S. Zrníc, 2005: The joint polarization experiment: Polarimetric rainfall measurements and hydrometeor classification. *Bull. Amer. Meteor. Soc.*, **86**, 809–824, <https://doi.org/10.1175/BAMS-86-6-809>.
- , M. Diederich, P. F. Zhang, and C. Simmer, 2014: Potential utilization of specific attenuation for rainfall estimation, mitigation of partial beam blockage, and radar networking. *J. Atmos. Oceanic Technol.*, **31**, 599–619, <https://doi.org/10.1175/JTECH-D-13-00038.1>.
- Schuurmans, J. M., F. C. van Geer, and M. F. P. Bierkens, 2011: Remotely sensed latent heat fluxes for model error diagnosis: A case study. *Hydrol. Earth Syst. Sci.*, **15**, 759–769, <https://doi.org/10.5194/hess-15-759-2011>.
- Seo, B. C., B. Dolan, W. F. Krajewski, S. A. Rutledge, and W. Petersen, 2015: Comparison of single- and dual-polarization-based rainfall estimates using NEXRAD data for the NASA Iowa Flood Studies project. *J. Hydrometeor.*, **16**, 1658–1675, <https://doi.org/10.1175/JHM-D-14-0169.1>.
- Shuttleworth, W. J., and J. S. Wallace, 1985: Evaporation from sparse crops—An energy combination theory. *Quart. J. Roy. Meteor. Soc.*, **111**, 839–855, <https://doi.org/10.1002/qj.49711146910>.
- Stisen, S., M. F. McCabe, J. C. Refsgaard, S. Lerer, and M. B. Butts, 2011: Model parameter analysis using remotely sensed pattern information in a multi-constraint framework. *J. Hydrol.*, **409**, 337–349, <https://doi.org/10.1016/j.jhydrol.2011.08.030>.
- , J. Koch, T. O. Sonnenborg, J. C. Refsgaard, S. Bircher, R. Ringgaard, and K. H. Jensen, 2018: Moving beyond runoff calibration—Multi-variable optimization of a surface-subsurface-atmosphere model. *Hydrol. Processes*, <https://doi.org/10.1002/hyp.13177>, in press.
- Villarini, G., J. A. Smith, M. L. Baecck, R. Vitolo, D. B. Stephenson, and W. F. Krajewski, 2011: On the frequency of heavy rainfall for the midwest of the United States. *J. Hydrol.*, **400**, 103–120, <https://doi.org/10.1016/j.jhydrol.2011.01.027>.
- Wang, Y. D., J. Zhang, A. V. Ryzhkov, and L. Tang, 2013: C-band polarimetric radar QPE based on specific differential propagation phase for extreme typhoon rainfall. *J. Atmos. Oceanic Technol.*, **30**, 1354–1370, <https://doi.org/10.1175/JTECH-D-12-00083.1>.
- Wealands, S. R., R. B. Grayson, and J. P. Walker, 2005: Quantitative comparison of spatial fields for hydrological model assessment—Some promising approaches. *Adv. Water Resour.*, **28**, 15–32, <https://doi.org/10.1016/j.advwatres.2004.10.001>.
- Yang, T., Q. X. Shao, Z. C. Hao, X. Chen, Z. X. Zhang, C. Y. Xu, and L. M. Sun, 2010: Regional frequency analysis and spatio-temporal pattern characterization of rainfall extremes in the Pearl River basin, China. *J. Hydrol.*, **380**, 386–405, <https://doi.org/10.1016/j.jhydrol.2009.11.013>.
- Younger, P. M., J. E. Freer, and K. J. Beven, 2009: Detecting the effects of spatial variability of rainfall on hydrological modelling within an uncertainty analysis framework. *Hydrol. Processes*, **23**, 1988–2003, <https://doi.org/10.1002/hyp.7341>.
- Zrníc, D. S., and A. Ryzhkov, 1996: Advantages of rain measurements using specific differential phase. *J. Atmos. Oceanic Technol.*, **13**, 454–464, [https://doi.org/10.1175/1520-0426\(1996\)013<0454:AORMUS>2.0.CO;2](https://doi.org/10.1175/1520-0426(1996)013<0454:AORMUS>2.0.CO;2).